Design of quasi-symplectic propagators for Langevin dynamics

Simone Melchionna

February 5, 2008

SOFT-INFM-CNR, Department of Physics,
University of Rome La Sapienza,
P.le A.Moro 2, 00185, Rome, Italy

Correspondence: simone.melchionna@roma1.infn.it

February 5, 2008

Abstract

A vector field splitting approach is discussed for the systematic derivation of numerical propagators for deterministic dynamics. Based on the formalism, a class of numerical integrators for Langevin dynamics are presented for single and multiple timestep algorithms.

1 Introduction

The design of efficient and stable propagators is a central issue in the numerical study of classical and quantum systems. The way the state is propagated in time affects the quality of the simulated trajectory and static and dynamical ensemble averages, as much as the capacity to efficiently sample the available phase space.

Molecular Dynamics (MD) is the reference technique for the study of condensed matter systems [1]. Although, in principle, time discretization introduces an error which affects the statistics of the simulated systems, from the point of view of numerical control, MD offers a number of advantages. The underlying conservative and time-reversible nature of the Hamiltonian equations allows to monitor the quality of the dynamics as a function of the time step and the implemented interaction potential. Despite energy conservation is never attained in practice, keeping the energy fluctuations small is a necessary, although not sufficient, condition to avoid statistical bias in the computed averages.

The lack of drifts in the energy has long been recognized as a key requirement to control long-time stability. In mechanical terms, long-time stability follows from the symplectic nature of the propagator, a distinguishing feature of Hamiltonian dynamics [2]. In statistical terms, symplecticity implies exact conservation of measure in phase space, i.e. the possibility of applying Gibbs statistical mechanics to a well-defined and conserved ensemble of physical states. Moreover, symplecticity and time-reversibility are the basis to combine MD with Monte Carlo, in the so-called Hybrid Monte Carlo method [3], where the acceptance test removes any systematic bias due to the usage of a large timestep. Another important benefit of conservative dynamics is the possibility to employ ad-hoc quasi-Hamiltonian dynamics to sample ensembles different from the microcanonical, a popular choice being the Nosè-Hoover dynamics [1] by retaining the same pleasant features of the energy conserving dynamics.

One of the most popular numerical propagators originates from the pioneering work of Verlet [4] who employed an algorithm first introduced by Störmer [5]. The distinguishing features of the Verlet (also known as Störmer-Verlet) algorithm are its simplicity, as compared to predictor-corrector or Runge-Kutta methods [2], time-reversibility and the lack of numerical drifts. Therefore, in spite of its limited accuracy, being quadratic in the timestep, the Verlet propagator stands as the reference algorithm for MD. Some years after its discovery, a systematic derivation based on an operatorial splitting approach was proposed [6]. By considering the case of a separable Hamiltonian, the so-called Trotter factorization [7] was applied to derive the Verlet algorithm and its equivalents, in particular the Velocity Verlet (VV) version. A crucial benefit of such operatorial approach is its generalization to a multiple time step (MTS) scheme [8], which extends the single time step (STS) one. In the MTS, the advance in time of the phase space state follows the temporal evolution of forces in an asyncronous way, i.e. by treating fast and slow interactions at different levels. The efficiency of simulation can thus be improved up to one order of magnitude in favourable circumstances.

The operatorial approach is very elegant and fruitful. However, it can be cumbersome in the systematic derivation of propagators in different contexts. A typical case is provided by non-separable Hamiltonians, finding wide application in the description in generalized coordinates or in presence of velocity-dependent forces (see e.g. [9]). In this case, symplectic and robust numerical schemes, such as the Generalized Leapfrog, have been known for some time [2]. Another wide class of dynamics is represented by stochastic equations of motion, such as the all-famous Langevin equation. In this context, the operatorial route has again been applied by considering the stochastic noise as a time-dependent perturbation [10] or by evolving the state according to a Fokker-Planck propagator [11, 12]. It has to be mentioned that for a first-order stochastic differential equation, a fourth-order accurate scheme has been derived [11] based on a Runge-Kutta propagation of the deterministic component. However, such scheme requires to compute high order derivatives, thus being rather complicated for practical applications in condensed matter, and is not specifically designed to reduce to symplectic in the limit of zero friction. In general, due to the presence of both stochastic and deterministic forces, conventional operatorial calculus should be applied with some care[13]. As a matter of fact, a previous operatorial approach was shown to overestimate the accuracy of the numerical trajectory [10, 14].

Stochastic dynamics applied to condensed matter systems has recently received renewed attention for several reasons: i) recent emphasis on multi-scale methods is based on the simultaneous evolution of atoms and surrounding hydrodynamic fields, the latter been treated via Lattice Boltzmann dynamics. In this technique noise serves to effectively enforce fluctuation-dissipation relations in otherwise decoupled systems, thus requiring

stable and accurate numerical integrators [15, 16]; ii) improving the stability of deterministic MTS algorithms to large timesteps can be achieved by attaching to each degree of freedom massive thermostats to overdamp the atomic motion [17] or by stabilizing the motion by an overdamped Langevin friction. The notion of Langevin stabilization is not new, and has already been proposed in the literature in different forms [18, 19]; iii) Langevin dynamics is an effective mean of attaching the system to a thermal reservoir, thus allowing to sample the canonical ensemble without the introduction of artificial dynamics with memory (e.g. Nosê-Hoover).

In the present paper we will consider a deterministic class of propagators, the Verlet being a particular case for separable Hamiltonian dynamics, derived from a different perspective, based on splitting the phase space vector field rather than by approximating the Liouvillean. The trajectory representation employed here is clearly equivalent to the operatorial one. However, in the operatorial approach one evaluates the involved time integrals at current times (generating so-called "forward" update) while in the trajectory representation one can evaluate such integrals with more general interpolation schemes. The present approach allows to establish a clear connection between the algorithm and eventual approximations supplementing the propagation of the trajectory. In the same spirit, the technique is extended to Langevin dynamics for both STS and MTS approaches. We will derive numerical integrators for Langevin dynamics accurate to second-order in the timestep and reducing to symplectic Verlet ones in the limit of zero friction. Our derivation treats momenta as natural elements of the algorithm, an aspect which has recently been debated [20]. Some numerical tests demonstrate the statistical consistency of the schemes for STS and MTS propagations. Moreover, the stochastic MTS scheme applied to a realistic system made of water molecules, therefore comprising intramolecular, excluded volume and electrostatic forces, is shown to achieve improved numerical stability and correct configurational statistics.

2 Deterministic dynamics

Let us consider the $6N$ phase space point $x = \{q_i, p_i\}_{i=1,3N}$ associated to the autonomous equations of motion written as

$$\dot{x} = SV(x) \tag{1}$$

where S is a generic matrix with constant elements and $V(x)$ a generic vector field. In case of Hamiltonian dynamics the vector field reads

$$SV(x) = \mathcal{S}\partial H(x) \tag{2}$$

where $\partial \equiv \{\frac{\partial}{\partial q_i}, \frac{\partial}{\partial p_i}\}$. $\mathcal{S}$ is the so-called symplectic matrix

$$\mathcal{S} = \begin{pmatrix} 0 & I_d \\ -I_d & 0 \end{pmatrix} \tag{3}$$

where the $3N \times 3N$ block matrices 0 and I_d are the null and identity matrix, respectively. Therefore, $\mathcal{S}$ has a trivial inverse $\mathcal{S}^{-1} = \mathcal{S}^T = -\mathcal{S}$, expressing the fact that the symplectic transformation is norm preserving ($\|\mathcal{S}g\| = \|g\|$). For separable Hamiltonians $V(x)$ reduces to the familiar form

$$\partial H(x) = \begin{pmatrix} -F(q) \\ p/m \end{pmatrix} \tag{4}$$

where F and p refer to the $3N$ components of forces and momenta and m are the masses, assumed to be equal for the sake of simplicity.

In order to integrate numerically eq.(1), we consider the generic decomposition

$$S = U + L \tag{5}$$

where, for Hamiltonian flow, S reduces to $\mathcal{S}$ and U and L specialize to

$$\mathcal{U} = \begin{pmatrix} 0 & I_d \\ 0 & 0 \end{pmatrix} \qquad \mathcal{L} = \begin{pmatrix} 0 & 0 \\ -I_d & 0 \end{pmatrix} \tag{6}$$

We employ now a splitting technique to integrate the equations of motion. The idea is to approximate the vector field (1) as the superposition of two linearly independent vector fields, each one associated to the equations of motion

$$\frac{dx}{d\lambda} = LV(x) \tag{7}$$

and

$$\frac{dx}{d\mu} = UV(x) \tag{8}$$

Symbolically, each evolution operator is briefly indicated as $\frac{d}{d\lambda}$ and $\frac{d}{d\mu}$ [21], where $\frac{d}{dt}$ refers to the complete evolution (1), associated to the updates

$$x_o \to e^{h\frac{d}{d\lambda}} x_o \tag{9}$$

and

$$x_o \to e^{h\frac{d}{d\mu}} x_o \tag{10}$$

together with the composite evolutions $x_o \to e^{h\frac{d}{d\mu}} e^{h\frac{d}{d\lambda}} x_o$ and $x_o \to e^{h\frac{d}{d\lambda}} e^{h\frac{d}{d\mu}} x_o$. The two operators $\frac{d}{d\lambda}$ and $\frac{d}{d\mu}$ do not commute in general, i.e. the so-called Lie bracket $[\frac{d}{d\lambda}, \frac{d}{d\mu}] \neq 0$, implying that the distance between the evoluted points has error of order h^2, i.e. $e^{h\frac{d}{d\lambda}} e^{h\frac{d}{d\mu}} x_o = (e^{h\frac{d}{d\mu}} e^{h\frac{d}{d\lambda}} + h^2[\frac{d}{d\lambda}, \frac{d}{d\mu}])x_o + O(h^3)$ [21]. On the other hand, by considering the symmetric composition $e^{\frac{h}{2}\frac{d}{d\lambda}} e^{\frac{h}{2}\frac{d}{d\mu}} e^{\frac{h}{2}\frac{d}{d\mu}} e^{\frac{h}{2}\frac{d}{d\lambda}}$ this approximates the true trajectory to second-order accuracy in a timestep h.

Associated with the two evolutions (9,10) we now consider either the exact solution of the dynamical equations for each vector field, $\phi_{L,h}(x_o)$ and $\phi_{U,h}(x_o)$, or the maps based on the four elementary approximations to the time integrals $L\int_0^h V(t)dt$ and $U\int_0^h V(t)dt$ as

$$\begin{aligned}
\Phi_{L,h}(x_o) &\equiv x_o + hLV_o \\
\bar{\Phi}_{L,h}(x_o) &\equiv x_o + hLV_h \\
\Phi_{U,h}(x_o) &\equiv x_o + hUV_o \\
\bar{\Phi}_{U,h}(x_o) &\equiv x_o + hUV_h
\end{aligned} \tag{11}$$

where the bar denotes evaluation of the integral at the upper extremum of the interval (requiring in principle the inversion of an implicit equation). Conversely, the operatorial route is always based on evaluations at the lower extremum, thus allowing for more flexibility in the design of numerical algorithms in the present treatment. Nevertheless, the vector field splitting and the approximated integrals are two independent sources of numerical error.

The Euler-A propagator is defined by the composition

$$x_{h/2} = \Phi_{U,h/2} \circ \bar{\Phi}_{L,h/2}(x_o) \tag{12}$$

and the Euler-B scheme is defined by the composition

$$x_{h/2} = \Phi_{L,h/2} \circ \bar{\Phi}_{U,h/2}(x_o) \tag{13}$$

It is easily shown that Euler-B is the adjoint of Euler-A, where the adjoint of the map M_h is defined as $M_h^\star = M_{-h}^{-1}$ [5]. A symmetrized algorithm can be constructed as

$$x_h = \Phi_{L,h/2} \circ \bar{\Phi}_{U,h/2} \circ \Phi_{U,h/2} \circ \bar{\Phi}_{L,h/2}(x_o) \tag{14}$$

or by interchanging $U \leftrightarrow L$. Given the presence of pairs of the type $\bar{\Phi}_L \circ \Phi_L$ and $\bar{\Phi}_U \circ \Phi_U$ in the consecutive applications of eq.(14), the global propagator is based on evaluation of the integrals via the midpoint method, being accurate to quadratic order in the timestep. The scheme is symmetric with respect to time-inversion, since

$$x_o = \bar{\Phi}_{L,-h/2} \circ \Phi_{U,-h/2} \circ \bar{\Phi}_{U,-h/2} \circ \Phi_{L,-h/2}(x_h) \tag{15}$$

For non-separable Hamiltonians, the propagator (14) reduces to the Generalized Leapfrog [2]. The updating scheme reads

$$\begin{aligned} p^\star &= p_o - \frac{h}{2}\frac{\partial H}{\partial q}(q_o, p^\star) \\ q_h &= q_o + \frac{h}{2}\left[\frac{\partial H}{\partial p}(q_h, p^\star) + \frac{\partial H}{\partial p}(q_o, p^\star)\right] \\ p_h &= p^\star - \frac{h}{2}\frac{\partial H}{\partial q}(q_h, p^\star) \end{aligned} \tag{16}$$

where $p^\star$ indicates intermediate values of the momenta.

For a separable Hamiltonian each update of the map $x_h = \phi_{\mathcal{L},h/2} \circ \phi_{\mathcal{U},h} \circ \phi_{\mathcal{L},h/2}(x_o)$ becomes explicit, the resulting scheme given by the celebrated Velocity Verlet (VV)

$$\begin{aligned} p^\star &= p_o + \frac{h}{2}F(q_o) \\ q_h &= q_o + h\frac{p^\star}{m} \\ p_h &= p^\star + \frac{h}{2}F(q_h) \end{aligned} \tag{17}$$

The so-called Position Verlet propagator (PV) is similarly obtained by composing the maps as $x_h = \phi_{\mathcal{U},h/2} \circ \phi_{\mathcal{L},h} \circ \phi_{\mathcal{U},h/2}(x_o)$.

A central property of the Generalized Leapfrog and Verlet-based propagators is symplecticity. Generally speaking, the symplectic property applies to Hamiltonian dynamics, which obeys the relation [2]

$$J^T \mathcal{S}^{-1} J = \mathcal{S}^{-1} \tag{18}$$

where J is the Jacobian of the transformation $x(0) \to x(t)$. This property is related to the fact that the dynamics is invariant under canonical transformations. Conservation of measure follows directly, since eq.(18) implies $\|J\|^2 = 1$.

An explicit calculation proves that the Euler-A scheme for Hamiltonian flow is symplectic. In fact,

$$J = \frac{\partial x_h}{\partial x_o} = I + h\mathcal{S}\partial^2 H_{h/2}\frac{\partial x_{h/2}}{\partial x_o} = I + h\mathcal{S}\Omega_{h/2}(I - h\mathcal{L}\Omega_{h/2})^{-1} = (I + h\mathcal{L}\Omega_{h/2})(I - h\mathcal{U}\Omega_{h/2})^{-1} \tag{19}$$

where $\Omega = \Omega^T \equiv \partial^2 H$ and having using the fact that $x_o = x_{h/2} - h\mathcal{L}\partial H_{h/2}$, so that $\frac{\partial x_{h/2}}{\partial x_o} = \left(\frac{\partial x_o}{\partial x_{h/2}}\right)^{-1} = (I - h\mathcal{L}\Omega_{h/2})^{-1}$. Finally,

$$\begin{aligned} J^T\mathcal{S}^{-1}J - \mathcal{S}^{-1} &= \left[I + \frac{h}{2}\mathcal{S}\Omega_{h/2}(I - \frac{h}{2}\mathcal{L}\Omega_{h/2})^{-1}\right]^T \mathcal{S}^{-1}\left[I + \frac{h}{2}\mathcal{S}\Omega_{h/2}(I - \frac{h}{2}\mathcal{L}\Omega_{h/2})^{-1}\right] - \mathcal{S}^{-1} \\ &= [(I - \frac{h}{2}\mathcal{L}\Omega_{h/2})^{-1}]^T\left[\frac{h}{2}\Omega_{h/2}\mathcal{L}\Omega_{h/2} + \frac{h}{2}\Omega_{h/2}\mathcal{U}\Omega_{h/2} + \frac{h}{2}\Omega_{h/2}\mathcal{S}^T\Omega_{h/2}\right] \\ &\quad \times(I - \frac{h}{2}\mathcal{L}\Omega_{h/2})^{-1} = 0 \end{aligned} \tag{20}$$

A central property of symplectic propagators regards the presence of a shadow Hamiltonian exactly conserved by the numerical flow and written as a power expansion in the timestep, $\tilde{H}(x) = H(x) + \sum_n h^n G^n(x)$, such that the associated vector field is equal to $\Delta x \equiv x_h - x_0 = hS\partial\tilde{H}$. An explicit expression for the leading $G^n(x)$ terms can be derived with leading terms $hG^1(x)$ and $-hG^1(x)$ for symplectic Euler-A and Euler-B schemes respectively. As a consequence, VV scheme has leading term of h^2 order [2]. Clearly, on the $\tilde{H} = const$ manifold, measure is preserved, since $\partial \cdot \Delta x = 0$, and one can define a shadow distribution function $\tilde{f}(y,t)$ such that a Liouville equation is exactly satisfied

$$\tilde{f}(y,h) - \tilde{f}(y,0) + \frac{1}{h}\Delta x\partial\tilde{f}(y,0) = 0 \tag{21}$$

3 Langevin dynamics

Let us consider the Langevin dynamics

$$\left(\begin{array}{c} \dot{q} \\ \dot{p} \end{array} \right) = \left(\begin{array}{c} p/m \\ F(q) - \gamma p + \eta(t) \end{array} \right) \tag{22}$$

where $\eta(t)$ is a white noise, with

$$\begin{aligned} \langle \eta(t) \rangle &= 0 \\ \langle \eta(t)\eta(t') \rangle &= g^2 \delta(t-t') \end{aligned} \tag{23}$$

γ is the friction coefficient, $g^2 \equiv 2\gamma mkT$, and the deterministic part arises from a separable Hamiltonian. The present treatment, applied to the case of a scalar constant friction, can be rapidly extended to the case of a position-dependent or tensorial friction, as used in the presence of hydrodynamics or confined diffusion. The equations of motion are written in compact notation as

$$\dot{x} = (\mathcal{S} + mG)\partial H(x) + \tilde{\eta}(t) \tag{24}$$

where

$$G = -\gamma \left(\begin{array}{cc} 0 & 0 \\ 0 & I_d \end{array} \right) \qquad \tilde{\eta}(t) = \left(\begin{array}{c} 0 \\ \eta(t) \end{array} \right) \tag{25}$$

Different algorithms can be employed to integrate the Langevin dynamics, the most popular being the Euler and the Heun ones [22] whereas, in presence of interatomic forces, different algorithms have been proposed in the past [23, 24, 25, 26, 27]. The numerical solution of (22) presents, in general, non trivial mixing between deterministic and stochastic terms. This is shown by a standard argument [28]. Let us consider an elementary step written in Euler form, $\delta x_t = f_o t + \tilde{W}_t$, where $f(x) = (S + mG)\partial H(x)$ and $\tilde{W}_t = \int_0^t \tilde{\eta}(s)ds$ is the Wiener increment [30]. The Euler step is exact at sufficiently small time t and can now be nested in the integral of (24), leading to the formal solution

$$\begin{aligned} x_h - x_0 &= \int_0^h dt(f_0 + \partial f_0 \delta x_t + \frac{1}{2}\partial^2 f_0 \delta x_t^2 + ... + \tilde{\eta}_t) = \\ &= f_0 h + \tilde{W}_h + \partial f_0 f_0 \frac{h^2}{2} + \partial f_0 \int_0^h dt \tilde{W}_t + O(h^{5/2}) \end{aligned} \tag{26}$$

This expression shows that, in order to obtain the desired accuracy, a second order integrator for the deterministic component is needed together with a proper handling of the stochastic forces. In fact, to second order, deterministic and stochastic terms mix up via a linear functional of the noise, $\partial f_0 \int_0^h dt \tilde{W}_t$, scaling as $h^{3/2}$. Due to the separable Hamiltonian, direct inspection reveals that this term reduces simply to $1/m \int_0^h dt W_t$ for positions and $-\gamma \int_0^h dt W_t$ for momenta, i.e. arising from inertial and Langevin forces only. In principle, the purely deterministic terms $(f_0 h + \partial f_0 f_0 \frac{h^2}{2} + ...)$, accounting for the effect of the operator $e^{hf\partial}(x_0)$, cannot be separated from the stochastic ones via the simplistic operatorial splitting, at the expense of missing the mixed term in eq.(26). However, it should be possible to decompose the vector fields by separating the evolution of the configurational degrees of freedom from that of the momenta, by considering the splitting

$$\dot{x} = (\mathcal{L} + mG)\partial H(x) + \tilde{\eta}(t) \tag{27}$$

and

$$\dot{x} = \mathcal{U}\partial H(x) \tag{28}$$

Eq. (27) has solution given by the exact map

$$\psi_{\mathcal{L},h}(x_o) \equiv \left(\begin{array}{c} q_0 \\ e^{-\gamma h} p_0 + \frac{(1-e^{-\gamma h})}{\gamma} F_0 + Q_h \end{array} \right) \tag{29}$$

where the following stochastic integral appears

$$Q_h = e^{-\gamma h} \int_o^h ds e^{\gamma s} \eta(s) = \mathcal{N}(mkT(1-e^{-2\gamma h})) \tag{30}$$

with $\mathcal{N}(\sigma^2)$ being a gaussian variable with zero mean and variance σ^2.

We define the stochastic Euler-A as

$$x_{h/2} = \phi_{\mathcal{U},h/2} \circ \psi_{\mathcal{L},h/2}(x_o) \tag{31}$$

and the stochastic Euler-B as

$$x_{h/2} = \psi_{\mathcal{L},h/2} \circ \phi_{\mathcal{U},h/2}(x_o) \tag{32}$$

By symmetric composition, the Stochastic Velocity Verlet (SVV) is constructed as

$$x_h = \psi_{\mathcal{L},h/2} \circ \phi_{\mathcal{U},h} \circ \psi_{\mathcal{L},h/2}(x_o) \tag{33}$$

where we have collapsed $\phi_{\mathcal{U},h/2} \circ \phi_{\mathcal{U},h/2} = \phi_{\mathcal{U},h}$. The complete updating scheme reads

$$\begin{aligned} p^\star &= e^{-\gamma h/2} p_o + \frac{(1-e^{-\gamma h/2})}{\gamma} F(q_o) + \mathcal{N}^{(1)}(mkT(1-e^{-\gamma h})) \\ q_h &= q_o + \frac{h}{m} p^\star \\ p_h &= e^{-\gamma h/2} p^\star + \frac{(1-e^{-\gamma h/2})}{\gamma} F(q_h) + \mathcal{N}^{(2)}(mkT(1-e^{-\gamma h})) \end{aligned} \tag{34}$$

where $\mathcal{N}^{(1)}$ and $\mathcal{N}^{(2)}$ are two independent realizations of the process (30) and thus, the algorithm requires two extractions of random numbers per timestep.

By reverse composition, the stochastic position Verlet (SPV) is constructed, reading

$$x_h = \phi_{\mathcal{U},h/2} \circ \psi_{\mathcal{L},h} \circ \phi_{\mathcal{U},h/2}(x_o) \tag{35}$$

where it is easily shown that $\psi_{\mathcal{L},h/2} \circ \psi_{\mathcal{L},h/2} = \psi_{\mathcal{L},h}$, having collapsed the sum of independent gaussian terms into a single gaussian extraction. In explicit terms, SPV reads

$$\begin{aligned} q^\star &= q_0 + \frac{h}{2m} p_0 \\ p_h &= e^{-\gamma h} p_o + \frac{(1-e^{-\gamma h})}{\gamma} F(q^\star) + \mathcal{N}(mkT(1-e^{-2\gamma h})) \\ q_h &= q^\star + \frac{h}{2m} p_h \end{aligned} \tag{36}$$

The accuracy of the SVV and SPV algorithms is evaluated below. At first, let us consider the explicit solution of eq.(22)

$$\begin{aligned} x_t &= \begin{pmatrix} q_0 + \int_0^t dt' e^{-\gamma t'} \frac{1}{m} \left[p_0 + \int_0^{t'} dt'' e^{\gamma t''} F(q(t'')) \right] + \frac{1}{m} \int_0^t dt' e^{-\gamma t'} \int_0^{t'} dt'' e^{\gamma t''} \eta(t'') \\ e^{-\gamma t} p_0 + e^{-\gamma t} \int_0^t dt' e^{\gamma t'} F(q(t')) + e^{-\gamma t} \int_0^t dt' e^{\gamma t''} \eta(t') \end{pmatrix} \\ &= \begin{pmatrix} q_0 + \left(\frac{1-e^{-\gamma t}}{\gamma m} \right) p_0 + \frac{1}{\gamma m} \int_0^t dt' F(q(t')) \left(1 - e^{\gamma (t-t')} \right) + \frac{1}{\gamma m} (W_t - Q_t) \\ e^{-\gamma t} p_0 + e^{-\gamma t} \int_0^t dt' e^{\gamma t'} F(q(t')) + Q_t \end{pmatrix} \end{aligned} \tag{37}$$

where we used the rule $\int_0^t dt' \int_0^{t'} dt'' = \int_0^t dt'' \int_{t''}^t dt'$. The variable Q_t is easily computed to give rise to the following covariances $\langle Q_t W_t \rangle = g^2(1-e^{-\gamma t})/\gamma$ and $\langle Q_t^2 \rangle = g^2(1-e^{-2\gamma t})/2\gamma$ [30]. As a result, for a constant

force the covariance matrix of positions and momenta reads

$$\begin{aligned}
\langle \Delta q_t^2 \rangle &= \frac{g^2}{m^2\gamma^3}\left(\gamma t - 2(1-e^{-\gamma t}) + \frac{(1-e^{-2\gamma t})}{2}\right) = \frac{g^2 t^3}{3m^2} + O(t^4) \\
\langle \Delta q_t \Delta p_t \rangle &= \frac{g^2}{2m\gamma^2}\left(1-e^{-\gamma t}\right)^2 = \frac{g^2 t^2}{2m}(1-\gamma t) + O(t^4) \\
\langle \Delta p_t^2 \rangle &= \frac{g^2}{2\gamma}\left(1-e^{-2\gamma t}\right)
\end{aligned} \tag{38}$$

We now confront these terms with the corresponding ones appearing in SVV and SPV. SVV generates the following noise terms

$$\begin{aligned}
\Delta q_h &= \frac{h}{m} Q_{h/2}^{(1)} \\
\Delta p_h &= e^{-\gamma h/2} Q_{h/2}^{(1)} + Q_{h/2}^{(2)}
\end{aligned} \tag{39}$$

with covariances given by $\langle \Delta q_h^2 \rangle = g^2 h^2 (1-e^{-\gamma h})/2\gamma m^2 = g^2 h^3/2m^2 + O(h^4)$, $\langle \Delta q_h \Delta p_h \rangle = g^2 h (e^{-\gamma h/2} - e^{-3\gamma h/2})/2\gamma m = g^2 h^2 (1-\gamma h)/2m + O(h^4)$ and $\langle \Delta p_h^2 \rangle = g^2 (1-e^{-2\gamma h})/2\gamma$. Similarly, for SPV we find that $\langle \Delta q_h^2 \rangle = g^2 h^3/4m + O(h^4)$, $\langle \Delta q_h \Delta p_h \rangle = g^2 h^2 (1-\gamma h)/2m + O(h^4)$ and $\langle \Delta p_h^2 \rangle = g^2 (1-e^{-2\gamma h})/2\gamma$. In conclusion, for both SVV and SPV positional variance coincides with (38) up to quadratic order, position-momentum covariance up to cubic order, and momentum to any order, of the form sought.

Before proceeding further, we wish to make a couple of comments regarding the derived schemes. The first is that in numerical applications, one is usually concerned with the usage of uniform random number generators in lieu of the expensive Guassian ones. This is typically possible for the simple Euler scheme [31] in which a uniform variate is used to sample a gaussian process up to second moment. In principle, a second order accuracy in the propagator would require the usage of a linear combination of uniform variates in order to sample the gaussian distribution up to the forth moment. In our treatment, we have used the information up to the second moment of the stochastic processes Q_t or W_t. This circumstance, mirrored by the presence of two random variables arise in the Verlet-like class of integrators, allows to employ uniform random number generators in the present case. The second comment regards the SVV scheme, whose integration of the deterministic part has already appeared in the literature [32], and shown to be equivalent to the one proposed by van Gunsteren and Berendsen [23]. In the present approach, however, the noise terms are not imposed to be equal to the matrix (38), as in previous approaches, but rather emerge spontaneously, ultimately due to the decomposition of the underlying vector field.

The derivation of SVV and SPV is based on the exact integration of eqs. (27,28). However, one could integrate the decomposed dynamics by evaluating the time integrals with expressions analogous to (11), but for the stochastic case. In particular, by defining the maps

$$\begin{aligned}
\Psi_{\mathcal{L},h}(x_o) &\equiv \begin{pmatrix} q_0 \\ -\gamma h p_0 + h F_0 + W_h \end{pmatrix} \\
\bar{\Psi}_{\mathcal{L},h}(x_o) &\equiv \begin{pmatrix} q_0 \\ -\gamma h p_h + h F_0 + W_h \end{pmatrix}
\end{aligned} \tag{40}$$

and the compositions $\bar{\Psi}_{\mathcal{L},h/2} \circ \phi_{\mathcal{U},h} \circ \Psi_{\mathcal{L},h/2}$ or $\phi_{\mathcal{U},h/2} \circ \bar{\Psi}_{\mathcal{L},h/2} \circ \Psi_{\mathcal{L},h/2} \circ \phi_{\mathcal{U},h/2}$ another pair of Verlet-like algorithms is derived. The updating scheme analogous to SVV, which we name SVVm, then reads

$$\begin{aligned}
p^\star &= \frac{1}{1+\gamma h/2}\left\{ p_o + \frac{h}{2} F(q_o) + \mathcal{N}^{(1)}(mkT\gamma h) \right\} \\
q_h &= q_o + \frac{h}{m} p^\star \\
p_h &= (1 - \frac{\gamma h}{2}) p^\star + \frac{h}{2} F(q_h) + \mathcal{N}^{(2)}(mkT\gamma h)
\end{aligned} \tag{41}$$

A straightforward calculation shows that the corresponding moments are again accurate to second order with the ones of eq.(38).

It is easily verified that the deterministic VV and PV propagators are recovered from SVV and SPV (or the approximate SVVm) in the limit $\gamma \to 0$. If morever the Jacobian is phase-space independent, the method is called quasi-symplectic [29]. Let us consider, as an example, the stochastic Euler-A method,

$$\frac{\partial x_o}{\partial x_{h/2}} = e^{-Gh/2} + \left(\frac{e^{\gamma h/2}-1}{\gamma}\right)\mathcal{L}\Omega_{h/2} \tag{42}$$

and

$$J = \frac{\partial x_h}{\partial x_o} = \left(\frac{\partial x_h}{\partial x_{h/2}}\right)\left(\frac{\partial x_o}{\partial x_{h/2}}\right)^{-1} = \left(I + \frac{h}{2}\mathcal{U}\Omega_{h/2}\right)\left(e^{-Gh/2} - \left(\frac{e^{\gamma h/2}-1}{\gamma}\right)\mathcal{L}\Omega_{h/2}\right)^{-1} \tag{43}$$

so that, for a block diagonal Ω, $\|J\| = e^{-3N\gamma h/2}$. The same result is derived for the stochastic Euler-B scheme. Globally, SVV and SPV have a Jacobian given by$\|J\| = e^{-3N\gamma h}$. Moreover, this property allows to employ the propagators in a Hybrid Monte Carlo scheme, based on an underlying Langevin dynamics.

The statistical distribution associated to the numerical Langevin flow map is readily derived. Its evolution is given by

$$\tilde{f}(y,t+h) - \tilde{f}(y,t) = \int dx \left[\delta(y - x_t - \Delta x) - \delta(y - x_t)\right] \tilde{f}(x_t,t) = \int dx \sum_{n=1}^{\infty} \frac{(-1)^n}{n!}\delta(y-x_t)\partial^n \left(\langle \Delta x^n \rangle_\eta \tilde{f}(x,t)\right) \tag{44}$$

having Taylor expanded the delta function and integrated by parts. Moreover, $\langle\cdot\rangle_\eta$ indicates averaging over noise. By considering a first order scheme, e.g. the stochastic Euler A, and up to a first order dependence in the timestep, we write $\Delta x = \Delta x_H + \Delta x_\gamma$, where Δx_H is the variation due to Hamiltonian dynamics (at $\gamma = 0$), being the vector field arising from the shadow Hamiltonian $\tilde{H}$ of the corresponding Hamiltonian numerical scheme. Moreover, Δx_γ arises from the dissipative and random terms.

By retaining the first two terms of the right hand side of eq. (44) we arrive at the following evolution equation

$$\tilde{f}(y,t+h) - \tilde{f}(y,t) = -\frac{1}{h}\langle \Delta x_H \rangle_\eta \partial \tilde{f}(y,t) - \partial\left\{G\left[\frac{1}{h}\langle \Delta x_\gamma \rangle_\eta + mk\tilde{T}\partial\right]\right\}\tilde{f}(y,t) \tag{45}$$

where we used the fact that, according to standard Langevin analysis, $\langle \Delta p^2 \rangle_\eta = mk\tilde{T}h$ which dominates over $\langle \Delta q^2 \rangle_\eta \sim h^2$ and $\tilde{T}$ is an effective numerical temperature associated to the momentum fluctuations. Moreover, in deriving (45) explicit use of the symplectic character of the Hamiltonian flow has been used. Equation (45) represents the numerical Liouville equation coupled to a Fokker-Planck evolution, with stationary solution $\tilde{f} \propto e^{-\tilde{H}/k\tilde{T}}$.

For a second order quasi-symplectic numerical scheme, the information gained from the elementary building blocks (Euler A and B) can be used and at equilibrium the Boltzmann distribution is recovered up to order h^2. In this case, the Boltzmann factor contains the shadow Hamiltonian of the underlying scheme at $\gamma = 0$.

We perform some numerical tests primarily aimed at controlling the quality of the integrators and the convergence of the first two moments to the theoretical expectations. An elementary test is provided by the one-dimensional stochastic oscillator, with potential $U(q) = q^2/2$, integrated in a single timestep scheme with the SVV and SPV algorithms. We use $h = 0.1$, $\gamma = 0.1$ and $T = 0.5$. The timestep is well below the stability limit $h = 2$, valid for the purely deterministic VV method. The computed momentum and configurational distributions, $P_{sim}(p,t)$ and $P_{sim}(q,t)$, should converge towards the theoretical forms, $P(p,\infty) \sim e^{-\beta p^2/2}$ and $P(q,\infty) \sim e^{-\beta U(q)}$, respectively. The rate of convergence of the normalized histograms is monitored by the norms

$$\begin{aligned} E_p(T) &= \alpha_p \max_{i\in N} |P_{sim}(p_i,T) - P(p_i,\infty)| \\ E_q(T) &= \alpha_q \max_{i\in N} |P_{sim}(q_i,T) - P(q_i,\infty)| \end{aligned} \tag{46}$$

where N is the number of bins and α_q and α_p are arbitrary factors, as a function of the sampling time window T. In Fig. 1 the histograms for the SVV motion are shown, and similar profiles are found for the SPV case.

The plot shows that the kinetic and configurational moments are correctly sampled down to the distant tails of the distributions. The rate of convergence (Fig. 2) is systematic, and faster for the configurational counterpart.

We next follow the momentum and position second moments (i.e. the so-called kinetic and configurational temperatures) in the high timestep/high friction limit. In this circumstance, it is often reported that Langevin integrators produce distinct and systematic departures of kinetic and configurational temperatures from the present input, such that equipartion is violated, i.e. $\langle q^2 \rangle \neq \langle p^2 \rangle \neq kT$ [33]. Here, we employ the information from the underlying Hamiltonian propagator to analyze such behavior and showing that, given the quasi-symplectic form of the integrator, very accurate results can be obtained as compared to non quasi-symplectic ones. The VV update reads

$$\begin{aligned} q_h &= (1 - \frac{h^2}{2})q_0 + hp_0 \\ p_h &= (1 - \frac{h^2}{2})p_0 - h(1 - \frac{h^2}{4})q_0 \end{aligned} \tag{47}$$

showing that, to second order, the dynamics arises from a shadow Hamiltonian of the form $\tilde{H} = \frac{1}{2}p^2 + \frac{1}{2}(1 - \frac{h^2}{4})q^2$. In spite of the symmetry of the equations of motion in the q, p variables, the fact that the discretized evolution presents a biased, reduced force constant is apparently odd. Moreover, if now the dynamics is equipped with the Langevin thermostat, it is understood that equipartition is violated, i.e. positional fluctuations deviate quadratically from the kinetic ones [33].

However, if we now consider the PV algorithm, constructed via a shift of the updating algorithm by half timestep with respect to the VV one, the update reads

$$\begin{aligned} q_h &= (1 - \frac{h^2}{2})q_0 + h(1 - \frac{h^2}{4})p_0 \\ p_h &= (1 - \frac{h^2}{2})p_0 - hq_0 \end{aligned} \tag{48}$$

showing that, by a simple shift of the "observation" time, the mass of the particle is rescaled by the same factor $(1 - h^2/4)$. Such factor is a manifestation of the symmetry of the original continuous dynamics, although it shows up at shifted times as compared to the VV case. If the Langevin thermostat is added, one observes a systematic shift of the kinetic temperature from the input one at increasing timesteps. The outcome of the present argument is that in order to obtain equilibrated observables, functions of position or velocity separately, one should samples these quantities at unequal times, representing optimal sampling points along the trajectory.

In Fig. 4 we plot the configurational and kinetic temperature of the harmonic oscillator sampled at mid ($(m + \frac{1}{2})h$, $m = 0, 1, 2...$) and full (mh, $m = 0, 1, 2...$) time steps. The data show that SVV produces excellent averages for both kinetic and configurational temperature at high timesteps, only if these are sampled at unequal times. It is important to notice that such averages are extremely robust, in fact they keep close to the theoretical value up to timestep $h = 2$. Raising friction has the effect of shifting configurational temperature at mid-steps by about 10%, while kinetic temperature at full-steps weakly deviates from the input value. In conclusion, at moderate friction one has a way to sample well equilibrated quantities since the ballistic behavior is well under control, while for increasing γ systematic errors appears in the fluctuations.

The accuracy of the SVV, SPV and SVVm integrators to sample the dynamical evolution is checked for the same stochastic harmonic oscillator for $kT = 1$ and $\gamma = 0.1$. Given the initial condition $q(0) = p(0) = 0$, the second moments are given by [34]

$$\begin{aligned} \langle q^2(t) \rangle - \langle q(t) \rangle^2 &= 1 + \frac{e^{-t}}{3}[-4 + \cos\sqrt{3}t - \sqrt{3}\sin\sqrt{3}t] \\ \langle p^2(t) \rangle - \langle p(t) \rangle^2 &= 1 + \frac{e^{-t}}{3}[-4 + \cos\sqrt{3}t + \sqrt{3}\sin\sqrt{3}t] \end{aligned} \tag{49}$$

In Fig. 3 we report the numerical relative error of the moments from the theoretical expectation (49) for times $\gg h$. The convergence of results towards the theoretical values is apparent and follows a h^2 dependence in all cases, qualifying the methods as weakly second order accurate.

4 Multiple time step dynamics

In this section, the multiple time step splitting extensions for both the deterministic and stochastic dynamics are derived, both with underlying separable Hamiltonian. We consider the decomposition of the Hamiltonian into slow and fast components as

$$\partial H(x) = \partial H^s(x) + \partial H^f(x) \tag{50}$$

The standard practice in MTS algorithms is to integrate the slow and fast components separately, with timesteps h and h/n, where n is a positive integer ($n > 1$), respectively. A convenient decomposition is given by

$$\partial H^s = \begin{pmatrix} -F^s(q) \\ 0 \end{pmatrix} \qquad \partial H^f = \begin{pmatrix} -F^f(q) \\ p/m \end{pmatrix} \tag{51}$$

where $F^s(q)$ and $F^s(q)$ are the slow and fast components of the interatomic forces. The MTS version of the VV algorithm is given by

$$x_h = \phi^s_{\mathcal{L},h/2} \circ \left[\phi^f_{\mathcal{L},h/2n} \circ \phi_{\mathcal{U},h/n} \circ \phi^f_{\mathcal{L},h/2n} \right]^n \circ \phi^s_{\mathcal{L},h/2}(x_o) \tag{52}$$

best known as the RESPA algorithm [8], which is again measure preserving and second order accurate in time. An equivalent MTS version of the position Verlet propagator can be constructed along similar lines.

The stochastic version of the MTS algorithm is obtained by associating the frictional and noise Langevin forces to the fast part of the mechanical forces. For example, the MTS version of the propagator generalizing the SVV method, reads

$$x_h = \phi^s_{\mathcal{L},h/2} \circ \left[\psi^f_{\mathcal{L},h/2n} \circ \phi_{\mathcal{U},h/n} \circ \psi^f_{\mathcal{L},h/2n} \right]^n \circ \phi^s_{\mathcal{L},h/2}(x_o) \tag{53}$$

We evaluate the performance of the MTS-SVV scheme by considering a one-dimensional model provided by a stochastic oscillator with potential energy $U(x) = 4.5x^2 + 0.025x^4$. The quartic term of the potential is associated to the slow forces. The timesteps are $\pi/3$ and $\pi/30$ for the slow and fast propagators respectively. Friction is set to 1.0 and 0.001 and temperature is set to 0.5. For this choice of the timesteps, the deterministic MTS-VV is known to produce a resonance phenomenon [35, 17], an effect that limits its applicability in more complex situations, such as in presence of intramolecular bond stretching motion.

Fig. 5 illustrates the convergence towards the expected probabilities of the configurational and momentum distributions monitored by the norms (46). The results exhibit a systematic convergence for all values of friction chosen. Despite its simplicity, the test demonstrates the consistency in the definition of momenta even in a MTS type of propagation.

As a more realistic test, we consider a system made of 256 water molecules and modelled by the flexible TIP3P force field (see ref.[9] for details on parameters). Following [36], the forces are splitted into four different levels, although other choices could be made [37]. The first, innermost level handles the fast intramolecular bond stretching and angular forces. The second, third and fourth levels integrate the short-range forces arising from the Lennard-Jones and direct-space Coulomb interactions (via the Ewald technique [1]) in cut-off regions of $[0, 5]$, $[5, 7.5]$ and $[7.5, 9.5]$ Å, respectively. Moreover, the fourth level handles the reciprocal space part of the Ewald interactions. Friction is set to $\gamma = 1\ ps^{-1}$ and $\gamma = 10^3\ ps^{-1}$. Temperature is set to $T = 300\ K$. The simulations are run with timesteps $(h_o, n_1 h_o, n_2 h_o, n_3 h_o)$ associated to the four sub-propagators, by setting $h_o = 0.25\ fs$ for the fast propagation and choosing for the integers (n_1, n_2, n_3) the values $(1, 1, 1)$ (STS), $(8, 1, 1)$, $(8, 2, 2)$ and $(8, 2, 4)$. All simulations are run for 20 ps total time.

In Fig. 6 the three radial distribution functions for the oxygen-oxygen, oxygen-hydrogen and hydrogen-hydrogen atoms are reported for the $(1, 1, 1)$ and the $(8, 2, 2)$ runs with $\gamma = 1\ ps^{-1}$. All data converge to the correct profiles, i.e. the $(1, 1, 1)$ run, and equal to a set of independent simulations made by using the purely deterministic Nosè-Hoover thermostat. However, the $(8, 2, 4)$ data present distorted profiles for $\gamma = 1\ ps^{-1}$. A closer inspection shows that the kinetic temperature of the system is larger than the input value by approximately 30%, probably due to the large timestep associated to the reciprocal term of electrostatics. However, by simply increasing friction to $\gamma = 10^2\ ps^{-1}$, the simulation temperature approaches the input value, with relative difference being less than 5%, basically removing the distortion in the radial profiles. It should be noticed that

with the increased friction we maintain the condition $\gamma h < 2$ for each level of integration, which ensures that the configurational temperature remains close to the kinetic one. In conclusion, some overdamping of the dynamics is capable of reducing spurious temperature shifts and configurational bais together with stabilizing the separate propagators of the MTS scheme. A separate study should be undertaken to investigate more systematically the application of the Langevin MTS approach in the simulation of condensed matter systems.

5 Conclusions

The present paper described the integration of Hamiltonian and Langevin dynamical equations guided by a symplectic decomposition of the underlying vector field in phase space. In the purely deterministic case, taken together with appropriate quadrature formulae, the scheme provided the basis for the Verlet and Leapfrog family of numerical propagators, which are second-order accurate and applicable to general deterministic (e.g. non-separable Hamiltonian) dynamics.

In presence of stochastic forces the approach maintains a similar splitting of the underlying Hamiltonian dynamics. The vector field route proved convenient for the analytical treatment and we derived numerical propagators which are weakly second order accurate. The correctness of the configurational and momentum averages and their fluctuations was demonstrated in a series of numerical tests. The deterministic and stochastic propagators are rapidly extended to multiple time step dynamics, as often employed in the simulation of condensed matter systems. In deterministic multiple time step algorithms the presence of resonance phenomena limits the stability range of the methodology. By studying a realistic system composed by water, and adopting a standard stochastic stabilization with large timesteps, we found that a time-discretization error, manifesting itself as a spurious shift in temperature, was reduced by employing a rather large value of the friction coefficient.

The performances of the proposed schemes to treat more complex stochastic equations, such as dissipative particle dynamics [38], where dissipative and random forces depend in a non trivial way on momenta and positions, will be described in a forthcoming paper.

References

[1] D. Frenkel and B. Smit, Understanding molecular simulation, Academic Press, London (1996).

[2] B. Leimkuhler and S. Reich, Simulating Hamiltonian dynamics, Cambridge University Press (2005).

[3] S. Duane, A.D. Kennedy, B.J. Pendleton and D. Roweth, Phys. Lett. B, 216, 195 (1987).

[4] L. Verlet, Phys. Rev. 98, 159 (1967).

[5] E. Hairer, C. Lubich, G. Wanner, Geometric numerical integration: structure preserving algorithms for ordinary differential equations, Springer, Berlin (2002).

[6] M. Suzuki, Phys. Lett. A, 387, 165 (1992).

[7] H.F. Trotter, Proc. Am. Math. Soc. 545, 10 (1959).

[8] M.E. Tuckerman, B.J. Berne and G.J. Martyna, J. Chem. Phys. 1990, 97 (1992).

[9] S. Melchionna, Mol. Phys. 3045, 104 (2006).

[10] A. Ricci and G. Ciccotti, Mol. Phys. 1927, 101 (2003).

[11] H.A. Forbert and S.A. Chin, Phys. Rev. E 63, 016703 (2000).

[12] G. De Fabritiis, M. Serrano, P. Español, P.V. Coveney, Physica A 429, 361 (2006).

[13] F. Kühnemund and M. Wacker, Proc. Am. Math. Soc., 3569, 129 (2001).

[14] G. Ciccotti, personal communication.

[15] P. Ahlrichs and B. Dünweg, Int. J. Mod. Phys. C, 1429, 9 (1998).

[16] M. Fyta, S. Melchionna, E. Kaxiras and S. Succi, MultiScale Modelling and Sim., 1156, 5 (2006).

[17] P. Minary, M.E. Tuckerman and G.J. Martyna, Phys. Rev. Lett. 150201, 93 (2004).

[18] Q. Ma and J.A. Izaguirre, Proceedings of the ACM Symposium on Applied Computing SAC 03, 178 (2003).

[19] J.A. Izaguirre, D.P. Catarello, J.M. Wozniak and R.D. Skeel, J. Chem. Phys., 2090, 114 (2001).

[20] W. Wang and R.D. Skeel, Mol. Phys. 2149, 101 (2003).

[21] B. Shutz, Geometrical methods of mathematical physics, Cambridge University Press (1980).

[22] P.E. Kloeden, E. Platen, Numerical solution of stochastic differential equations, Springer, Berlin (1992).

[23] W.F. van Gunsteren and H.J.C. Berendsen, Mol. Phys. 637, 45 (1982).

[24] A. Brunger, C.L. Brooks and M. Karplus, Chem. Phys. Lett. 105, 637 (1984).

[25] M.P. Allen and D.J. Tildesley, Computer simulation of liquids, Clarendon press, Oxford (1987).

[26] R.D. Skeel and J.A. Izaguirre, Mol. Phys. 3885, 100 (2002).

[27] R. Mannella, Phys. Rev. E 041107, 69 (2004).

[28] R.L. Honeycutt, Phys. Rev. A, 600, 45 (1992).

[29] G.N. Milnstein and M.V. Tretyakov, "Numerical methods for Langevin type equations based on symplectic integrators", Preprint 727, Weierstrass Institut für Angewandte Analysis und Stochastik, Berlin (2002).

[30] C.W. Gardiner, Handbook of stochastic methods, Springer Verlag, Berlin (1985).

[31] B. Dünweg and W. Paul, Int. J. Mod. Phys. C, 817, 2 (1991).

[32] M.P. Allen, Mol.Phys., 599, 3 (1982).

[33] M.P. Allen, J. Phys. Chem. B, 3823, 110 (2006).

[34] S. Chadrasekhar, Rev. Mod. Phys., 1, 15 (1943).

[35] S.A. Chin, J. Chem. Phys. 8, 120 (2004).

[36] M. Marchi and P. Procacci, J. Chem. Phys., 5194, 109 (1998).

[37] X. Qian and T. Schlick, J. Chem. Phys. 5971, 116 (2002).

[38] P.J. Hoogerbrugge, J.M.V.A. Koelman, Europhys. Lett. **19**, 155 (1992). J.M.V.A. Koelman, P.J. Hoogerbrugge, Europhys. Lett. **21**, 363 (1993)

Captions

Figure 1: Histograms $P(q,T)$ and $P(p,T)$ in log-linear scale as a function of $z = q^2$ and $z = p^2$, respectively. Results generated with SVV for $T = 10^7 h$ simulation time. The x-axis has been rescaled by an arbitrary factor. Squares and circles refer to momentum and position distributions, respectively.

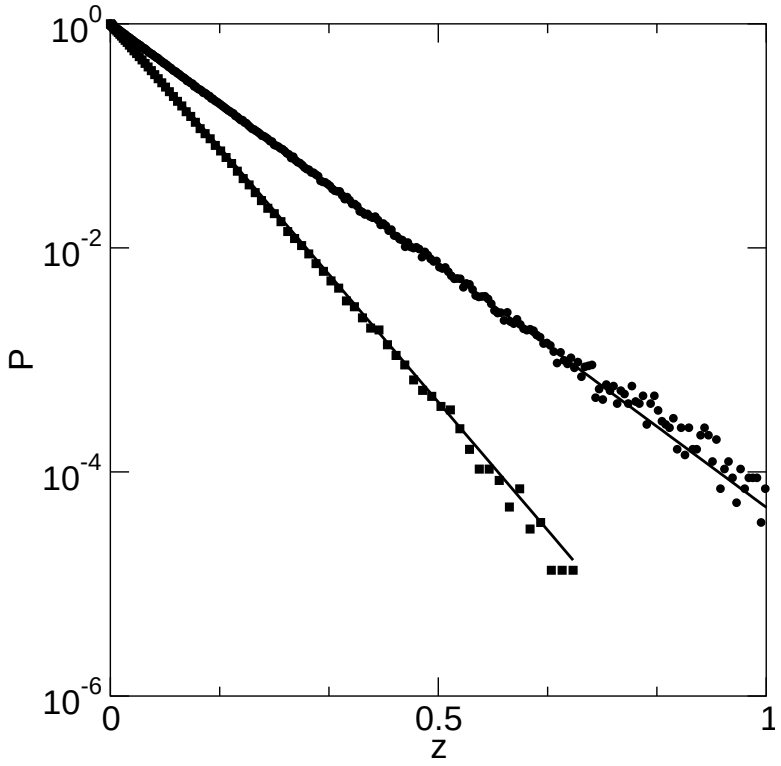
Figure 2: Norms E_p and E_q as a function of the sampling time window for the SVV algorithm. The solid line refers to E_q and the dashed line refers to E_p.

Figure 3: Relative error of second moments from the theoretical expectation (49) as a function of the timestep sampled for $t = 1$ and $q(0) = p(0) = 0$. Sampling time is 10^8 for each timestep chosen. Symbols refer to SVV data for positions (open circles) and momenta (filled circles), SPV data for positions (open squares) and momenta (filled squares), SVVm data for positions (open diamonds) and momenta (filled diamonds).The dashed line corresponds to h^2 and the dot-dashed line to h trends.

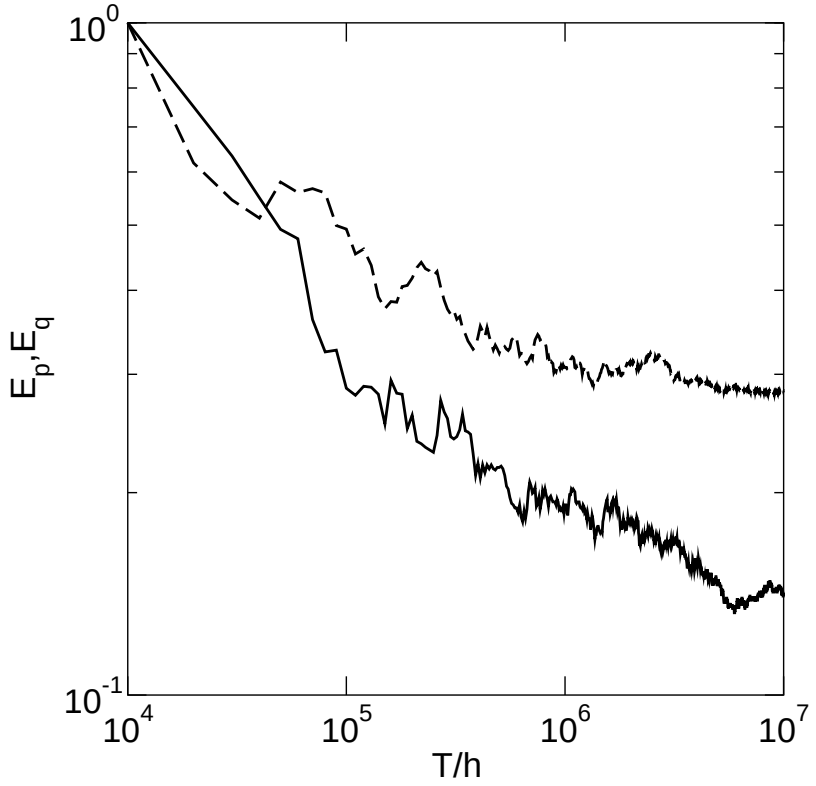
Figure 4: Deviation of configurational and kinetic temperatures from the input temperature for the stochastic harmonic oscillator propagated with SVV at $kT = 1$. Symbols: configurational data sampled at mid-step (circles), and full-step (stars), and kinetic data at mid-step (squares) and full-step (triangles). Filled symbols refer to $\gamma = 0.1$ and open symbols to $\gamma = 1.0$.

Figure 5: Convergence towards the expected distribution for the SVV algorithm in the single (thick lines) and multiple (thin lines) timestep formulations. Solid and dashed lines refer to configurational and momentum data, respectively.

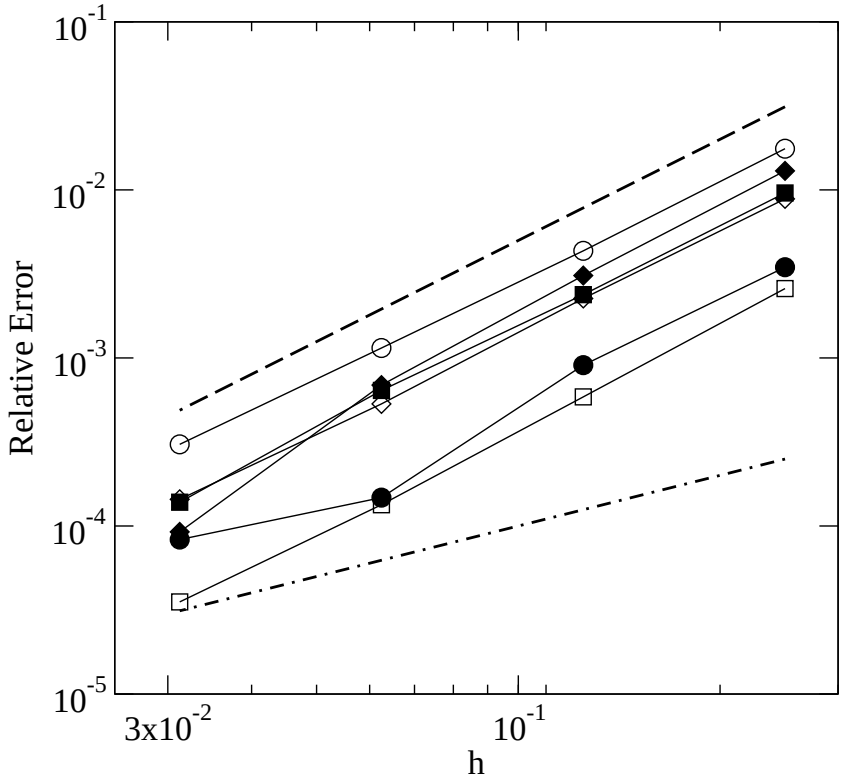
Figure 6: Radial distribution functions for water for the (1,1,1) (solid curves) and (8,2,2) (dashed curves) simulations with the MTS-SVV integrator (see text for details). The curves refer to oxygen-oxygen, oxygen-hydrogen and hydrogen-hydrogen profiles.



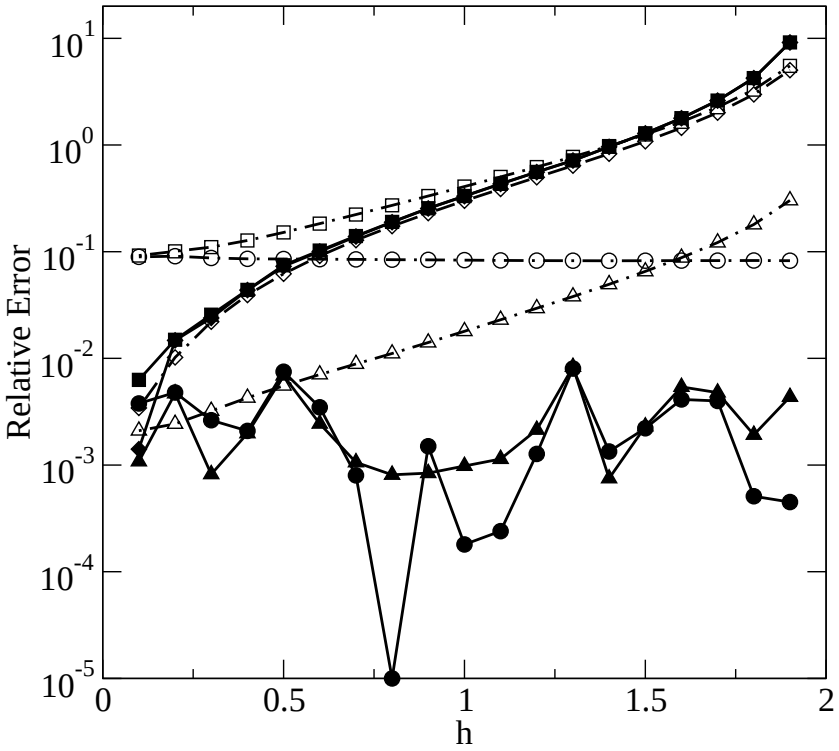
Melchionna, Fig.1



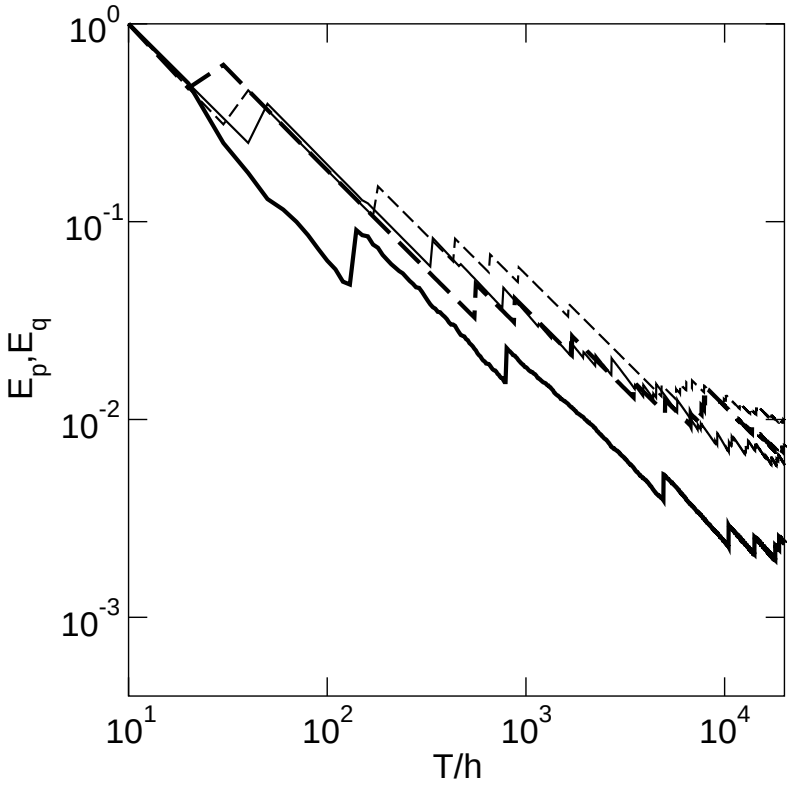
Melchionna, Fig.2



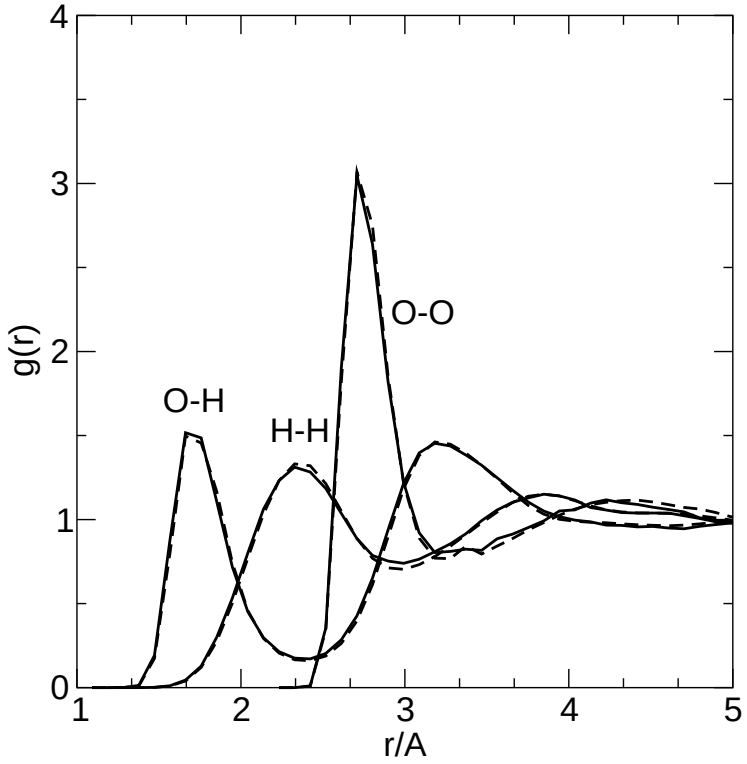
Melchionna, Fig.3



Melchionna, Fig.4



Melchionna, Fig. 5



Melchionna, Fig.6